確率的ファジィ意志決定について(最適化理論と数理構造)

岩本, 誠一

数理解析研究所講究録
第 864 巻 1994 年 203-213

確率的ファジィ意志決定について

岩本 誠一 (Seiichi IWAMOTO)

九州大学 経済学部 経済工学科

1 はじめに

『ファジィ環境下の意志決定』は R. Bellman and L. Zadeh, "Decision-making in a fuzzy environment", Management Science 17(1970), [3],141-164 ではじめて導入され，その後の議論に大きな影響を与えている ([2], [4], [6], [7], [8])．Bellman and Zadeh のアプローチはファジィ環境下とはいえ，基本的には動的計画法を用いてある最小型評価基準を最大化する意志決定過程を構成することであると考えられる．この最小型評価基準は所定の制約を満たしながらゴール（目標）に到達する迄の直積（履歴）空間におけるメンバーシップ関数である．特に，確定的なシステムでは最小型評価そのものを最大化している．また，確率的な推移システムでは最小型評価基準の期待値を最大化していると考えられる．

動的計画本来の立場からすれば，確定的システム にしろ確率的システムにしろ，動的計画法を用いた逐次最適化の解はなんらかの方法（例えば，列挙法）による同時最適化の解と一致するべきであると考えられる．この点，確定的なシステムにおける動的計画法の彼らの援用では確かに二つの方法による最適解は一致している．

しかし，確率的なシステムにおける動的計画法によるの彼らの最適解は列挙法による最適解とは一致していない．従って，確率的システムに対しては Bellman and Zadeh による動的計画法の適用は（少なくとも数学的に）理論的整合性に欠けると考えられる．

本論文では，確率的推移システムに対して Bellman and Zadeh とは異なる動的計画法を導入して,「逐次最適化＝同時最適化」を保証する最適解を与える．その基本的な考えは不変埋没原理 (Principle of Invariant Imbedding) である．具体的には，最小型基準の期待値最大化問題を新しいパラメータ λ を含む問題群に埋め込んで，そこで再帰式（第 1 の再帰式）を導き，これを解いて $\lambda = 1$ のとき，所与の問題の最適解が得られる．また，Bellman and Zadeh の再帰式における「逐次最適化$\neq$同時最適化」を彼らの数値例で検証する．

以下では 特に断わらない限り Bellman and Zadeh の記号を用いることにする．

2 確率的多段決定過程

本節では, Bellman and Zadeh [3],§ 4 ,§ 5 の記号・用語を用いる．この節では, まず§ 5 の確率的意志決定過程を一部引用してその内容を再検討する．

> 前節の（確定的）問題におけると同様にして, 終端時刻 N は固定され, 初期状態 x_0 が与えられているとする．システムの状態推移はマルコフ型条件付き確率 $p(x_{t+1}|x_t, u_t)$ で記述されるとする．このとき, 問題はファジィ制約 $C^0, \cdots, C^{N-1}$ を満たしながら, 時刻 N においてファジィゴールに到達する確率を最大にすることである．

もしファジィゴールG^NがXにおけるファジィ事象とみなされるならば,

$$\mathrm{Prob}(G^N|x_{N-1},u_{N-1}) = E\mu_{G^N}(x_N) = \sum_{x_N} p(x_N|x_{N-1},u_{N-1})\mu_{G^N}(x_N) \qquad (1)$$

で表される. ただし,E は条件付き期待値を表し, μ_{G^N} は所与のファジィゴールのメンバーシップ関数である.

ここで, 期待値の記号 $E\mu_{G^N}(x_N)$ の代わりにむしろ条件付き期待値を表わす記号 $E\mu_{G^N}(\cdot\,|x_{N-1},u_{N-1})$ を用いた方がよいことに注意しよう. このとき, 式 (1) は

$$\mathrm{Cond.Exp}(G^N|x_{N-1},u_{N-1}) = E\mu_{G^N}(\cdot\,|x_{N-1},u_{N-1}) = \sum_{x_N} p(x_N|x_{N-1},u_{N-1})\mu_{G^N}(x_N) \quad (2)$$

になる.

さて, 式 (1) は $\mathrm{Prob}(G^N|x_{N-1},u_{N-1})$ すなわち x_{N-1}, u_{N-1} の関数 $E\mu_{G^N}(x_N)$ を表していることに注意しょう. これは前述の (確定的) 問題においても $\mu_{G^N}(x_N)$ が確定的運動方程式

$$x_{t+1} = f(x_t,u_t), \quad t = 0,1,2,\cdots \qquad (3)$$

を通じて x_{N-1}, u_{N-1} で表されいることと同様である. したがって, $E\mu_{G^N}(x_N)$ を非確率的な場合における $\mu_{G^N}(x_N)$ と同じように取り扱えて, (いま考えている) 確率的な問題の解き方を (前述の) 確定的なそれに帰着することができる.

厳密には, 確定的再帰式

$$\mu_{G^{N-\nu}}(x_{N-\nu}) = \operatorname*{Max}_{u_{N-\nu}}(\mu_{N-\nu}(u_{N-\nu}) \wedge \mu_{G^{N-\nu+1}}(x_{N-\nu+1})) \qquad (4)$$

$$x_{N-\nu+1} = f(x_{N-\nu},u_{N-\nu}), \quad \nu = 1,\cdots,N, \qquad (5)$$

を確率的再帰式

$$\mu_{G^{N-\nu}}(x_{N-\nu}) = \operatorname*{Max}_{u_{N-\nu}}(\mu_{N-\nu}(u_{N-\nu}), E\mu_{G^{N-\nu+1}}(x_{N-\nu+1})) \qquad (6)$$

$$E\mu_{G^{N-\nu+1}}(x_{N-\nu+1}) = \sum_{x_{N-\nu+1}} p(x_{N-\nu+1}|x_{N-\nu},u_{N-\nu})\mu_{G^{N-\nu+1}}(x_{N-\nu+1}) \qquad (7)$$

に置き換えることができる. ただし, $\mu_{G^{N-\nu}}(x_{N-\nu})$ は, 時刻 $t = N-\nu+1$ におけるファジィゴールによって導かれた $t = N-\nu$ でのファジィゴールのメンバーシップを表している. ここに $\nu = 1,2,\cdots,N$.

ここで, 式 (6),(7) は正しくは

$$\mu_{G^{N-\nu}}(x_{N-\nu}) = \operatorname*{Max}_{u_{N-\nu}}[\mu_{N-\nu}(u_{N-\nu}) \wedge E\mu_{G^{N-\nu+1}}(\cdot\,|x_{N-\nu},u_{N-\nu})] \qquad (8)$$

$$E\mu_{G^{N-\nu+1}}(\cdot\,|x_{N-\nu},u_{N-\nu}) = \sum_{x_{N-\nu+1}} p(x_{N-\nu+1}|x_{N-\nu},u_{N-\nu})\mu_{G^{N-\nu+1}}(x_{N-\nu+1}) \qquad (9)$$

に置き換えた方が自然である. 事実, [3, pp. B154-B155] の例は式 (8),(9) によって計算されている ([4, pp.153], [6, pp.172], [7, pp.235], [8, pp.147] 参照).

次に, 条件付き期待値最適化問題:

$$\begin{array}{lll} \text{Maximize} & E[\,\mu_0(u_0)\wedge\mu_1(u_1)\wedge\cdots\wedge\mu_{N-1}(u_{N-1})\wedge\mu_{G^N}(x_N)\,] & \\ \text{subject to} & \text{(i)}_\text{n}\ \ x_{n+1}\sim p(\cdot\,|\,x_n,u_n) \quad 0\le n\le N-1 & (10)\\ & \text{(ii)}_\text{n}\ \ u_n\in U \quad 0\le n\le N-1. & \end{array}$$

を考える. ただし, E は, マルコフ型条件付き確率 $p(x_{n+1}|\,x_n,u_n)$, 政策 $\pi=\{\pi_0,\pi_1,...,\pi_{N-1}\}$, および初期状態 x_0 で定まる直積（履歴）空間 $U\times X\times U\times X\cdots\times U\times X$ 上の期待値（積分）作用素である. さて, 任意に与えられた $x_{N-\nu}$ に対して部分問題：

$$\mu_{G^{N-\nu}}(x_{N-\nu})=\text{Max}\ \ E[\,\mu_{N-\nu}(u_{N-\nu})\wedge\cdots\wedge\mu_{N-1}(u_{N-1})\wedge\mu_{G^N}(x_N) \\ |\,\text{(i)}_\text{m},\ \text{(ii)}_\text{m}\quad N-\nu\le m\le N-1\,] \qquad (11)$$

を考えよう. このとき, 従来の動的計画法の立場からすれば, 値 $\mu_{G^{N-\nu}}(x_{N-\nu})$ と関数 $\{\mu_{G^{N-\nu+1}}(x_{N-\nu+1})\}$ の間の関係式導きたいのである. しかし, この再帰関係式を導くことは幾分無理がある [5]. 従来の方法と異なった別のアプローチが必要になってくる. ここでは, この問題を新しいパラメータを含む問題群に埋め込み, 不変埋没原理を用いる. さて, 任意に与えられた状態 $x_{N-\nu}$ と区間 [0, 1] 上の実数 λ に対して部分最適化問題:

$$\mu_{G^{N-\nu}}(x_{N-\nu};\lambda)=\text{Max}\ \ E[\,\lambda\wedge\mu_{N-\nu}(u_{N-\nu})\wedge\cdots\wedge\mu_{N-1}(u_{N-1})\wedge\mu_{G^N}(x_N) \\ |\,\text{(i)}_\text{m},\ \text{(ii)}_\text{m}\quad N-\nu\le m\le N-1\,] \qquad (12) \\ 1\le\nu\le N$$

$$\mu_{G^N}(x_N;\lambda)=\lambda\wedge\mu_{G^N}(x_N) \qquad 0\le\lambda\le 1 \qquad (13)$$

を考えると, 次の再帰式が成立する:

定理 1

$$\mu_{G^{N-\nu}}(x_{N-\nu};\lambda)=\underset{u_{N-\nu}}{\text{Max}}\sum_{x_{N-\nu+1}}\mu_{G^{N-\nu+1}}(x_{N-\nu+1};\lambda\wedge\mu_{N-\nu}(u_{N-\nu})) \\ \times p(x_{N-\nu+1}|x_{N-\nu},u_{N-\nu}) \qquad (14) \\ x_{N-\nu}\in X,\quad 0\le\lambda\le 1 \qquad \nu=1,2,\cdots,N$$

$$\mu_{G^N}(x_N;\lambda)=\lambda\wedge\mu_{G^N}(x_N) \qquad x_N\in X,\quad 0\le\lambda\le 1. \qquad (15)$$

さて, 式 (14) の最大値に到達する $u_{N-\nu}$ の（任意の）値を $\tilde{\pi}_{N-\nu}(x_{N-\nu};\lambda)$ で表す. 列 $\tilde{\pi}=\{\tilde{\pi}_0,\tilde{\pi}_1,...,\tilde{\pi}_{N-1}\}$ をパラメータ化された問題 (12),(13) の **最適政策** という. 以下では, 1 変数関数 $\mu_{G^{N-\nu}}(x_{N-\nu})$ と 2 変数関数 $\mu_{G^{N-\nu}}(x_{N-\nu};\lambda)$ を区別することが重要である. 一般に, 両者は一致しない:

$$\mu_{G^{N-\nu}}(x_{N-\nu};\lambda)\neq\lambda\wedge\mu_{G^{N-\nu}}(x_{N-\nu}) \qquad \nu=1,2,\cdots,N-1. \qquad (16)$$

([5] を参照). しかし, λ の十分大きい値 $\hat{\lambda}$ に対しては等式:

$$\mu_{G^{N-\nu}}(x_{N-\nu})=\mu_{G^{N-\nu}}(x_{N-\nu};\hat{\lambda}) \qquad (17)$$

が成り立つ. たとえば, $\hat{\lambda}$ を

$$\hat{\lambda}\ge\text{Max}\ \ [\,\mu_{N-\nu}(u_{N-\nu})\wedge\cdots\wedge\mu_{N-1}(u_{N-1})\wedge\mu_{G^N}(x_N) \\ |\,\text{(i)}_\text{m},\ \text{(ii)}_\text{m}\quad N-\nu\le m\le N-1\,] \qquad (18)$$

を満たすように選べば，等式 (17) は成立する．このように，求める最大期待値 $\mu_{G^0}(x_0)$ はλ の十分大きい値 $\hat{\lambda}$ (≤ 1) に対する $\mu_{G^0}(x_0;\hat{\lambda})$ で与えられる：

$$\mu_{G^0}(x_0) = \mu_{G^0}(x_0;\hat{\lambda}). \tag{19}$$

尤も，十分大きいと言っても $\hat{\lambda}$ は $\hat{\lambda}=1$ でよい．それはメンバーシップが常に $0 \leq \mu_A(x) \leq 1$ で，その最小演算値と$\hat{\lambda}$ との最小比較であるからである．

さらに，確定的システムを確率的システムの特別なケースとみなすことによって，埋没問題の 2 変数間の確率的再帰式から Bellman and Zadeh の確定的再帰式 (4),(5):

$$\mu_{G^{N-\nu}}(x_{N-\nu}) = \operatorname*{Max}_{u_{N-\nu}} (\mu_{N-\nu}(u_{N-\nu}) \wedge \mu_{G^{N-\nu+1}}(x_{N-\nu+1})) \tag{20}$$

$$x_{N-\nu+1} = f(x_{N-\nu}, u_{N-\nu}), \quad \nu = 1, \cdots, N. \tag{21}$$

を演繹的に導くことができる．しかし，逆に確定的再帰式の単なる類推（Bellman and Zadeh の意味でのアナロジー）で確率的システムの結果を導くのは危険であり，理論的であるとは言い難い．

3 Bellman and Zadeh の例

本節では，逐次最適化（すなわち，我々の不変埋没原理に基づく動的計画法による最適化）が本来の求める同時最適化に一致することを確かめるために，Bellman and Zadeh の例 [3, pp. B154] を用いる．彼らのデータは次の通りである：

$$\mu_{G^2}(\sigma_1) = 0.3, \quad \mu_{G^2}(\sigma_2) = 1.0, \quad \mu_{G^2}(\sigma_3) = 0.8 \tag{22}$$

$$\mu_1(\alpha_1) = 1.0, \qquad \mu_1(\alpha_2) = 0.6 \tag{23}$$

$$\mu_0(\alpha_1) = 0.7, \qquad \mu_0(\alpha_2) = 1.0 \tag{24}$$

$u_t = \alpha_1$

$x_t \diagdown x_{t+1}$	σ_1	σ_2	σ_3
σ_1	0.8	0.1	0.1
σ_2	0.0	0.1	0.9
σ_3	0.8	0.1	0.1

$u_t = \alpha_2$

$x_t \diagdown x_{t+1}$	σ_1	σ_2	σ_3
σ_1	0.1	0.9	0.0
σ_2	0.8	0.1	0.1
σ_3	0.1	0.0	0.9

3.1 埋没問題に対する再帰式

本小節では，パラメータ λ を含んだ我々の再帰式：

$$\mu_{G^{N-\nu}}(x_{N-\nu};\lambda) = \operatorname*{Max}_{u_{N-\nu}} \sum_{x_{N-\nu+1}} \mu_{G^{N-\nu+1}}(x_{N-\nu+1}; \lambda \wedge \mu_{N-\nu}(u_{N-\nu}))$$
$$\times p(x_{N-\nu+1}|x_{N-\nu}, u_{N-\nu}) \tag{25}$$
$$\nu = 1, 2, \cdots, N$$

$$\mu_{G^N}(x_N;\lambda) = \lambda \wedge \mu_{G^N}(x_N) \qquad 0 \leq \lambda \leq 1 \tag{26}$$

に実際に Bellman and Zadeh の数値例をあてはめて，これを解く．
先ず，

$$N=2,\quad \mu_{G^2}(\sigma_1)=0.3,\quad \mu_{G^2}(\sigma_2)=1,\quad \mu_{G^2}(\sigma_3)=0.8,$$

とすると，

$$\begin{aligned}\mu_{G^2}(\sigma_1;\lambda)&=\lambda\wedge 0.3\\ \mu_{G^2}(\sigma_2;\lambda)&=\lambda\wedge 1\\ \mu_{G^2}(\sigma_3;\lambda)&=\lambda\wedge 0.8\end{aligned}\tag{27}$$

が得られる．次に，再帰式

$$\mu_{G^1}(x_1;\lambda)=\operatorname*{Max}_{u_1\in\{\alpha_1,\alpha_2\}}\sum_{x_2\in\{\sigma_1,\sigma_2,\sigma_3\}}\mu_{G^2}(x_2;\lambda\wedge\mu_1(u_1))p(x_2|x_1,u_1)\tag{28}$$

を $x_1=\sigma_1$ にあてはめると，

$$\mu_{G^1}(\sigma_1;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ 0.9\lambda+0.03 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ 0.57 & 0.6\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_1(\sigma_1;\lambda)=\begin{cases}\alpha_1,\ \ \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_2 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ \alpha_2 & 0.6\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

になることがわかる．$x_1=\sigma_2$, $x_1=\sigma_3$ に対しても同様にすると，次が得られる:

$$\mu_{G^1}(\sigma_2;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \lambda & 0.3\le\lambda\le 0.8 \quad \text{のとき}\\ 0.1\lambda+0.72 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_1(\sigma_2;\lambda)=\begin{cases}\alpha_1,\ \ \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_1 & 0.3\le\lambda\le 0.8 \quad \text{のとき}\\ \alpha_1 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\mu_{G^1}(\sigma_3;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ 0.9\lambda+0.3 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ 0.57 & 0.6\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_1(\sigma_3;\lambda)=\begin{cases}\alpha_1,\ \ \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_2 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ \alpha_2 & 0.6\le\lambda\le 1 \quad \text{のとき}.\end{cases}$$

最後の再帰式

$$\mu_{G^0}(x_0;\lambda)=\operatorname*{Max}_{u_0\in\{\alpha_1,\alpha_2\}}\sum_{x_1\in\{\sigma_1,\sigma_2,\sigma_3\}}\mu_{G^1}(x_1;\lambda\wedge\mu_0(u_0))p(x_1|x_0,u_0)\tag{29}$$

を解くと，

$$\mu_{G^0}(\sigma_1;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ 0.99\lambda+0.003 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ 0.9\lambda+0.057 & 0.6\le\lambda\le 0.8 \quad \text{のとき}\\ 0.09\lambda+0.705 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_0(\sigma_1;\lambda)=\begin{cases}\alpha_1,\quad \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_2 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ \alpha_2 & 0.6\le\lambda\le 0.8 \quad \text{のとき}\\ \alpha_2 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\mu_{G^0}(\sigma_2;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ 0.91\lambda+0.027 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ 0.1\lambda+0.513 & 0.6\le\lambda\le 0.7 \quad \text{のとき}\\ 0.1\lambda+0.513 & 0.7\le\lambda\le 0.8 \quad \text{のとき}\\ 0.01\lambda+0.585 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_0(\sigma_2;\lambda)=\begin{cases}\alpha_1,\quad \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_1,\quad \alpha_2 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ \alpha_1,\quad \alpha_2 & 0.6\le\lambda\le 0.7 \quad \text{のとき}\\ \alpha_2 & 0.7\le\lambda\le 0.8 \quad \text{のとき}\\ \alpha_2 & 0.8\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\mu_{G^0}(\sigma_3;\lambda)=\begin{cases}\lambda & 0\le\lambda\le 0.3 \quad \text{のとき}\\ 0.91\lambda+0.027 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ 0.1\lambda+0.513 & 0.6\le\lambda\le 0.7 \quad \text{のとき}\\ 0.583 & 0.7\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

$$\tilde{\pi}_0(\sigma_3;\lambda)=\begin{cases}\alpha_1,\quad \alpha_2 & 0\le\lambda\le 0.3 \quad \text{のとき}\\ \alpha_1 & 0.3\le\lambda\le 0.6 \quad \text{のとき}\\ \alpha_1 & 0.6\le\lambda\le 0.7 \quad \text{のとき}\\ \alpha_1 & 0.7\le\lambda\le 1 \quad \text{のとき}\end{cases}$$

が得られる. したがって, このときの条件付き最適化問題:

$$\begin{array}{ll}\text{Maximize} & E[\,\mu_0(u_0)\wedge\mu_1(u_1)\wedge\mu_{G^2}(x_2)\,]\\ \text{subject to} & (\mathrm{i})_{\mathrm{n}}\quad x_{n+1}\sim p(\cdot\mid x_n,u_n)\qquad n=0,1\\ & (\mathrm{ii})_{\mathrm{n}}\quad u_n\in\{\alpha_1,\alpha_2\}\qquad n=0,1\end{array}\tag{30}$$

は次の最大期待値を持つ:

$$\begin{aligned}\mu_{G^0}(\sigma_1)&=\mu_{G^0}(\sigma_1;1)=0.795\\ \mu_{G^0}(\sigma_2)&=\mu_{G^0}(\sigma_2;1)=0.595\\ \mu_{G^0}(\sigma_3)&=\mu_{G^0}(\sigma_3;1)=0.583.\end{aligned}\tag{31}$$

初期状態 $(x_0;1)$ からの最大期待値とそれを与える**最適行動**(最適履歴とも言う) は $(x_0;1)$ から最適政策 $\tilde{\pi}=\{\tilde{\pi}_0,\tilde{\pi}_1\}$ を用いることによって得られる. 図 1 は最適政策 $\tilde{\pi}$ からの最適行動のみならず, 対応する最大期待値を示している. もちろん, ここでは**行動**とは状態, 決定, 各段利得, 1 段推移確率から成る 4 項目の (時刻 0 , 1 , 2 までの) 3 段にわたる列である.

Figure 1: $(\sigma_1;1)$ からの最適行動と最大期待値

$(x_0;1.0)$	u_0	μ_0	p_0	$(x_1;\lambda_1)$	u_1	μ_1	p_1	$(x_2;\lambda_2)$	μ_2	経路確率	最小値	積
							0.1	$(\sigma_1;0.6)$	0.3	*0.01*	0.3	0.003
			0.1	$(\sigma_1;1.0)$	α_2	0.6	*0.9*	$(\sigma_2;0.6)$	1.0	*0.09*	0.6	0.054
							0.0	$(\sigma_3;0.6)$	0.8	*0.00*	0.6	0.000
							0.0	$(\sigma_1;1.0)$	0.3	*0.00*	0.3	0.000
$(\sigma_1;1.0)$	α_2	1.0	*0.9*	$(\sigma_2;1.0)$	α_1	1.0	*0.1*	$(\sigma_2;1.0)$	1.0	*0.09*	1.0	0.090
							0.9	$(\sigma_3;1.0)$	0.8	*0.81*	0.8	0.648
							0.1	$(\sigma_1;0.6)$	0.3	*0.00*	0.3	0.000
			0.0	$(\sigma_3;1.0)$	α_2	0.6	*0.0*	$(\sigma_2;0.6)$	1.0	*0.00*	0.6	0.000
							0.9	$(\sigma_3;0.6)$	0.8	*0.00*	0.6	0.000
										最大総期待値		**0.795**

図1では次のように記号を簡略化して用いている:

$$u_0 = \tilde{\pi}_0(x_0; 1),\quad \mu_0 = \mu_0(u_0),\quad p_0 = p(x_1 \mid x_0, u_0),\quad x_1 \sim p(\cdot \mid x_0, u_0),\quad \lambda_1 = 1 \wedge \mu_0$$
$$u_1 = \tilde{\pi}_1(x_1; \lambda_1),\quad \mu_1 = \mu_1(u_1),\quad p_1 = p(x_2 \mid x_1, u_1),\quad x_2 \sim p(\cdot \mid x_1, u_1),\quad \lambda_2 = \lambda_1 \wedge \mu_1$$
$$\mu_2 = \mu_{G^2}(x_2),\quad 最小値 = \mu_0 \wedge \mu_1 \wedge \mu_2,\quad 経路確率 = p_0 p_1,\quad 積 = 経路確率 \times 最小値$$

このように不変埋没原理を用いた動的計画法による最大期待値と最適行動は, また直接(すべての場合を調べ尽くす全数)列挙法によっても求めることができる. 次の表1はこの列挙法での(すべての場合も含む型で)最適解を初期状態毎にそれぞれ示している.

表1では次のようにやはり記号を簡略化して用いている.

$$履歴 = x_0\ \ u_0\ \ \mu_0(u_0)\ \ p(x_1 \mid x_0, u_0)\ \ x_1\ \ u_1\ \ \mu_1(u_1)\ \ p(x_2 \mid x_1, u_1)\ \ x_2$$
$$終端 = 終端利得 = \mu_{G^2}(x_2)$$
$$経路 = 経路確率 = p(x_1 \mid x_0, u_0) \times p(x_2 \mid x_1, u_1)$$
$$最小 = 最小値 = \mu_0(u_0) \wedge \mu_1(u_1) \wedge \mu_{G^2}(x_2)$$
$$積 = 経路 \times 最小 = 経路確率 \times 最小値$$
$$部分 = 部分期待値,\quad 総期 = 総期待値$$

さらに, 斜体数字は確率を表し, ゴチック体数字は上項と下項の2つうち大きい(最大の)期待値を選択していることを示している.

さて, 図1と表1をそれぞれ比較してみると, 初期状態 $x_1 = \sigma_1$ での最適解が図と表でそれぞれ一致していることがわかる. すなはち, 我々の不変埋没原理を用いた動的計画法による逐次最適解(図1)がいわゆる列挙法による同時最適解(表1)に一致していることがわかった. ところが次の小節で判明するように, Bellman and Zadeh の(いわゆる加法型利得系に対する従来の意味での)動的計画法による逐次最適解はこの列挙法による同時最適解に一致していない. 動的計画法本来の考え方であるところの, 所与の問題をこれを含むより広い1題群に埋め込んで, そこで再帰式を樹立する必要がある. Bellman and Zadeh の動的計画法では, 所与の問題を解くには十分に広い題群に埋め込んだことにはならない. 我々の動的計画法では, 所与の問題を解くには(利得系にパラメータλを導入した)1次元広い問題群に埋め込めば, 再帰式も導けて, しかもそれを完全に解くことができることがわかった. この問題では1次元広い問題群に埋め込んでうまくいったが, しかし一般には, 所与の問題をどれ位い大きな問題群に埋め込めば再帰式が導けて解けるかわからない. どのような問題群に埋め込むかがまさしく動的計画法そのものの問題と言えよう.

履歴									終端	経路	最小	積	部分	総期
σ_1	α_1	0.7	*0.8*	σ_1	α_1	1.0	*0.8*	σ_1	0.3	*0.64*	0.3	0.192		
σ_1	α_1	0.7	*0.8*	σ_1	α_1	1.0	*0.1*	σ_2	1.0	*0.08*	0.7	0.056	0.304	
σ_1	α_1	0.7	*0.8*	σ_1	α_1	1.0	*0.1*	σ_3	0.8	*0.08*	0.7	0.056		
σ_1	α_1	0.7	*0.8*	σ_1	α_2	0.6	*0.1*	σ_1	0.3	*0.08*	0.3	0.024		
σ_1	α_1	0.7	*0.8*	σ_1	α_2	0.6	*0.9*	σ_2	1.0	*0.72*	0.6	0.432	**0.456**	
σ_1	α_1	0.7	*0.8*	σ_1	α_2	0.6	*0.0*	σ_3	0.8	*0.0*	0.6	0		
σ_1	α_1	0.7	*0.1*	σ_2	α_1	1.0	*0.0*	σ_1	0.3	*0.0*	0.3	0		
σ_1	α_1	0.7	*0.1*	σ_2	α_1	1.0	*0.1*	σ_2	1.0	*0.01*	0.7	0.007	**0.070**	
σ_1	α_1	0.7	*0.1*	σ_2	α_1	1.0	*0.9*	σ_3	0.8	*0.09*	0.7	0.063		0.583
σ_1	α_1	0.7	*0.1*	σ_2	α_2	0.6	*0.8*	σ_1	0.3	*0.08*	0.3	0.024		
σ_1	α_1	0.7	*0.1*	σ_2	α_2	0.6	*0.1*	σ_2	1.0	*0.01*	0.6	0.006	0.036	
σ_1	α_1	0.7	*0.1*	σ_2	α_2	0.6	*0.1*	σ_3	0.8	*0.01*	0.6	0.006		
σ_1	α_1	0.7	*0.1*	σ_3	α_1	1.0	*0.8*	σ_1	0.3	*0.08*	0.3	0.024		
σ_1	α_1	0.7	*0.1*	σ_3	α_1	1.0	*0.1*	σ_2	1.0	*0.01*	0.7	0.007	0.038	
σ_1	α_1	0.7	*0.1*	σ_3	α_1	1.0	*0.1*	σ_3	0.8	*0.01*	0.7	0.007		
σ_1	α_1	0.7	*0.1*	σ_3	α_2	0.6	*0.1*	σ_1	0.3	*0.01*	0.3	0.003		
σ_1	α_1	0.7	*0.1*	σ_3	α_2	0.6	*0.0*	σ_2	1.0	*0.0*	0.6	0	**0.057**	
σ_1	α_1	0.7	*0.1*	σ_3	α_2	0.6	*0.9*	σ_3	0.8	*0.09*	0.6	0.054		
σ_1	α_2	1.0	*0.1*	σ_1	α_1	1.0	*0.8*	σ_1	0.3	*0.08*	0.3	0.024		
σ_1	α_2	1.0	*0.1*	σ_1	α_1	1.0	*0.1*	σ_2	1.0	*0.01*	1.0	0.01	0.042	
σ_1	α_2	1.0	*0.1*	σ_1	α_1	1.0	*0.1*	σ_3	0.8	*0.01*	0.8	0.008		
σ_1	α_2	1.0	*0.1*	σ_1	α_2	0.6	*0.1*	σ_1	0.3	*0.01*	0.3	0.003		
σ_1	α_2	1.0	*0.1*	σ_1	α_2	0.6	*0.9*	σ_2	1.0	*0.09*	0.6	0.054	**0.057**	
σ_1	α_2	1.0	*0.1*	σ_1	α_2	0.6	*0.0*	σ_3	0.8	*0.0*	0.6	0		
σ_1	α_2	1.0	*0.9*	σ_2	α_1	1.0	*0.0*	σ_1	0.3	*0.0*	0.3	0		
σ_1	α_2	1.0	*0.9*	σ_2	α_1	1.0	*0.1*	σ_2	1.0	*0.09*	1.0	0.09	**0.738**	
σ_1	α_2	1.0	*0.9*	σ_2	α_1	1.0	*0.9*	σ_3	0.8	*0.81*	0.8	0.648		**0.795**
σ_1	α_2	1.0	*0.9*	σ_2	α_2	0.6	*0.8*	σ_1	0.3	*0.72*	0.3	0.216		
σ_1	α_2	1.0	*0.9*	σ_2	α_2	0.6	*0.1*	σ_2	1.0	*0.09*	0.6	0.054	0.324	
σ_1	α_2	1.0	*0.9*	σ_2	α_2	0.6	*0.1*	σ_3	0.8	*0.09*	0.6	0.054		
σ_1	α_2	1.0	*0.0*	σ_3	α_1	1.0	*0.8*	σ_1	0.3	*0.0*	0.3	0		
σ_1	α_2	1.0	*0.0*	σ_3	α_1	1.0	*0.1*	σ_2	1.0	*0.0*	1.0	0	**0**	
σ_1	α_2	1.0	*0.0*	σ_3	α_1	1.0	*0.1*	σ_3	0.8	*0.0*	0.8	0		
σ_1	α_2	1.0	*0.0*	σ_3	α_2	0.6	*0.1*	σ_1	0.3	*0.0*	0.3	0		
σ_1	α_2	1.0	*0.0*	σ_3	α_2	0.6	*0.0*	σ_2	1.0	*0.0*	0.6	0	**0**	
σ_1	α_2	1.0	*0.0*	σ_3	α_2	0.6	*0.9*	σ_3	0.8	*0.0*	0.6	0		

表 1: 初期状態 σ_1 からの全行動と最大分枝（最適行動）の選択

3.2 Bellman and Zadeh の再帰式

本小節では, Bellman and Zadeh の数値データを用いて彼等のアプローチ [3] を再び検証する. 彼等は彼等の再帰式 (8),(9):

$$\mu_{G^{N-\nu}}(x_{N-\nu}) = \underset{u_{N-\nu}}{\mathrm{Max}}\left[\mu_{N-\nu}(u_{N-\nu}) \wedge E\mu_{G^{N-\nu+1}}(\cdot\ |x_{N-\nu}, u_{N-\nu})\right] \tag{32}$$

$$E\mu_{G^{N-\nu+1}}(\cdot\ |x_{N-\nu},u_{N-\nu}) = \sum_{x_{N-\nu+1}} p(x_{N-\nu+1}|x_{N-\nu},u_{N-\nu})\mu_{G^{N-\nu+1}}(x_{N-\nu+1}) \tag{33}$$

を適用している．しかし，このアプローチは，次に示されるいわば後ろ向きに各段毎に条件付き確率をとって得られた“確定的”逐次最適化問題を解いている:

$$\begin{array}{lll} \text{Maximize} & [\,\mu_0(\pi_0)\wedge E^{\pi_0}\,[\mu_1(\pi_1)\wedge E^{\pi_1}\mu_{G^2}(x_2)\,]\,] & \\ \text{subject to} & (\mathrm{i})_{\mathrm{n}}\quad x_{n+1}\sim p(\cdot|\ x_n,u_n) & n=0,1 \\ & (\mathrm{ii})_{\mathrm{n}}\quad \pi_n(x_n)\in\{\alpha_1,\alpha_2\} & n=0,1 \end{array} \tag{34}$$

ただし

$$\begin{aligned} E^{\pi_1}k(x_2) &= \sum_{x_2} k(x_2)p(x_2|x_1,\pi_1(x_1)) \qquad k=k(x_2)\ \text{のとき} \\ \mu_1(\pi_1) &= \mu_1(\pi_1(x_1)) \end{aligned}$$

はともに x_1 の関数で，

$$\begin{aligned} E^{\pi_0}l(x_1) &= \sum_{x_1} l(x_1)p(x_1|x_0,\pi_0(x_0)) \qquad l=l(x_1)\ \text{のとき} \\ \mu_0(\pi_0) &= \mu_0(\pi_0(x_0)) \end{aligned}$$

も x_0 の関数である．この（確定的になった）問題に対しては，いわゆる（第 3 節での）動的計画法を用いることができる．したがって，このとき π_0,π_1 の 2 変数同時最適化が 2 段逐次最適化に帰着できて，等式:

$$\begin{aligned} & \underset{\pi_0,\pi_1}{\text{Max}}\ [\,\mu_0(\pi_0)\wedge E^{\pi_0}\,[\mu_1(\pi_1)\wedge E^{\pi_1}\mu_{G^2}(x_2)\,]\,] \\ =\ & \underset{\pi_0}{\text{Max}}\,[\,\mu_0(\pi_0)\wedge E^{\pi_0}\ \underset{\pi_1}{\text{Max}}\,[\mu_1(\pi_1)\wedge E^{\pi_1}\mu_{G^2}(x_2)\,]\,] \end{aligned} \tag{35}$$

が成立する．この等式を再帰式で表すと，他でもなく Bellman and Zadeh の再帰式:

$$\mu_{G^1}(x_1) = \underset{u_1\in\{\alpha_1,\alpha_2\}}{\text{Max}}[\,\mu_1(u_1)\wedge \sum_{x_2\in\{\sigma_1,\sigma_2,\sigma_3\}} \mu_{G^2}(x_2)p(x_2|x_1,u_1)\,] \tag{36}$$

$$\mu_{G^0}(x_0) = \underset{u_0\in\{\alpha_1,\alpha_2\}}{\text{Max}}[\,\mu_0(u_0)\wedge \sum_{x_1\in\{\sigma_1,\sigma_2,\sigma_3\}} \mu_{G^1}(x_1)p(x_1|x_0,u_0)\,] \tag{37}$$

になる．彼等は前述の彼等のデータを用いてこの再帰式を後ろ向きに解いて，最適解:

$$\mu_{G^1}(\sigma_1)=0.6,\quad \mu_{G^1}(\sigma_2)=0.82,\quad \mu_{G^1}(\sigma_3)=0.6 \tag{38}$$

$$\pi_1(\sigma_1)=\alpha_1,\quad \pi_1(\sigma_2)=\alpha_1,\quad \pi_1(\sigma_3)=\alpha_2 \tag{39}$$

$$\mu_{G^0}(\sigma_1)=0.8,\quad \mu_{G^0}(\sigma_2)=0.62,\quad \mu_{G^0}(\sigma_3)=0.62 \tag{40}$$

$$\pi_0(\sigma_1)=\alpha_1,\quad \pi_0(\sigma_2)=\alpha_1,\ \alpha_2,\quad \pi_0(\sigma_3)=\alpha_1 \tag{41}$$

を得ている [3, pp.B154-B155]([7, pp.235-240] も参照)．しかし，$\mu_{G^0}(x_0)$, $\pi_0(x_0)$ の厳密な値は次のようになる:

$$\mu_{G^0}(\sigma_1)=0.798,\quad \mu_{G^0}(\sigma_2)=0.622,\quad \mu_{G^0}(\sigma_3)=0.622 \tag{42}$$

$$\pi_0(\sigma_1) = \alpha_2, \quad \pi_0(\sigma_2) = \alpha_1,\ \alpha_2, \quad \pi_0(\sigma_3) = \alpha_1. \tag{43}$$

この点, Bellman and Zadeh の計算式の適用もさることながら, 計算結果そのものにも正確でない所がある.

さて, 我々の確率的問題 (30) の最適解 (31) と Bellman and Zadeh の "確定的"問題 (34) のそれを比較検討してみよう. すると, 2つの問題 (30),(34) は同値でないことがわかる (詳細は [5] 参照):

$$\begin{aligned} &\operatorname*{Max}_{\pi}\ E^{\pi}[\,\mu_0(u_0)\wedge\mu_1(u_1)\wedge\mu_{G^2}(x_2)\,] \\ \neq\ &\operatorname*{Max}_{\pi_0}[\,\mu_0(\pi_0)\wedge E^{\pi_0}\operatorname*{Max}_{\pi_1}[\mu_1(\pi_1)\wedge E^{\pi_1}\mu_{G^2}(x_2)\,]\,]. \end{aligned} \tag{44}$$

前の第4.1節で示したように, パラメータ λ を用いた不変埋没原理による方法が前述の〈元来は同時最適化である〉問題 (30) を正確に（もっと厳密には, 同時最適化＝逐次最適化を保証するという意味で）逐次的に解いていることがわかる.

一般に, 2つの解は一般には一致しない. 数学的には定数 λ, 関数 g, 確率測度 p に対して, 加法系のとき等式

$$\int_X [\lambda + g(x)]dp(x) = \lambda + \int_X g(x)dp(x)$$

が成り立つが, 最小系のときは一般に

$$\int_X [\lambda \wedge g(x)]dp(x) \neq \lambda \wedge \int_X g(x)dp(x)$$

であるからである. 前者は期待値（積分）作用素の線形性に他ならないが, 後者は期待値作用素の一つの非線形性を示している.

非線形問題を線形問題と同様に（ないしは 線形問題のアナロジーとして）取り扱うわけにはいかない. この逆は往々にして可能である. 非線形問題には非線形特有のアプローチが必要であろう. 本論文では, それは不変埋没原理であった.

References

[1] Bellman, R.E.: *Dynamic Programming*, Princeton Univ. Press, NJ, 1957.

[2] Baldwin, J.F. and Pilsworth, B.W.: Dynamic programming for fuzzy systems with fuzzy environment, *Journal of Mathematical Analysis and Applications*, Vol.85 (1982), 1-23.

[3] Bellman, R.E. and Zadeh, L.A.: Decision-making in a fuzzy environment, *Management Science*, Vol.17 (1970), B141-B164.

[4] Esogbue, A.O. and Bellman, R.E.: Fuzzy dynamic programming and its extensions, *TIMS/Studies in the Management Sciences*, Vol.20 (1984), 147-167.

[5] Iwamoto, S.: Associative dynamic programs, under consideration.

[6] Kacprzyk, J.: Decision-making in a fuzzy environment with fuzzy termination time, *Fuzzy Sets and Systems*, Vol.1 (1978), 169-179.

[7] 水本 雅晴: ファジィ理論とその応用, サイエンス社, 1988.

[8] 小田中 敏男: 確率制御過程, 森北出版, 1976.